

Solution Concepts for Distributed Decision-Making without Coordination

Final Report for Grant WAG-1-03059

Peter A. Beling (PI) and Stephen D. Patek (co-PI)
Coordinated Systems Research Group
Department of Systems and Information Engineering
University of Virginia
{beling, patek}@virginia.edu

Abstract

Consider a single-stage problem in which we have a group N agents who are attempting to minimize the expected cost of their joint actions, without the benefit of communication or a pre-established protocol but with complete knowledge of the expected cost of any joint set of actions for the group. We call this situation a static coordination problem. The central issue in defining an appropriate solution concept for static coordination problems is considering how to deal with the fact that if the agents are faced with a set of multiple (mixed) strategies that are equally attractive in terms of cost, a failure of coordination may lead to an expected cost value that is worse than that of any of the strategies in the set. In this proposal, we describe the notion of a general coordination problem, describe initial efforts at developing a solution concept for static coordination problems, and then outline a research agenda that centers on activities that will be basis for obtaining a complete understanding of solutions to static coordination problems.

1 Overview

This document serves as the final report for the research conducted under grant NAG-1-03059. This research centers on the development of a solution concept for a problem in distributed decision making that may be posed as follows: There are M agents, each of which has an associated set of control actions that it may take. Each agent knows the actions available to every agent. All of the agents must simultaneously choose an action, and the M -tuple of chosen actions determines systems performance in a way that is known precisely to each of the agents before the actions are selected. The agents may not communicate information about the actions that they will take, though it is known by all that each will seek to maximize system performance.

There are three primary features that make this problem a special or limiting case of the general problem of distributed decision-making:

1. The agents share a common objective of maximizing some measure of system performance.
2. The agents each know how to predict system performance with certainty, given the actions selected by all the agents. Equivalently, we may say that the agents share a common model of system performance.
3. No communication between agents to coordinate actions is allowed.

This problem poses a significant, non-computational challenge only when there exists more than one set of control actions that optimizes systems performance.

In this report, we formulate the notion of a general coordination problem and then describe the solution concept for static coordination problems that constitutes a primary outcome of the research under grant NAG-1-03059.

2 Problem Statement

To motivate the notion of coordination processes (as fully developed in following sections), let us consider a simple example involving two people (Al and Betty) engaged in a telephone conversation. Suppose that in the middle of the conversation the line is cut and that both Al and Betty wish to resume their conversation as quickly as possible, but who should call whom? If both participants employ the same amount of time either picking up and dialing or waiting for the return call, then there might be a problem:

1. If both Al and Betty redial, then both will receive busy signals and the line will not be reconnected.
2. If both Al and Betty wait for the return call, then again the line will not be reconnected.
3. Only if one party or the other places the call will the connection be made.

Thus, unless Al and Betty established in advance a protocol for who must call back, they must play out a sequence of actions (dialing and/or waiting) for the connection to be remade. What complicates the problem is that Al and Betty do not have the opportunity to communicate with one another in order to establish the best way to proceed; the inability to communicate is the essence of the problem. We refer to the dynamic process of reestablishing their connection as a coordination process.

Even though the Dial-Wait problem is a very simple example, it exposes a number of interesting issues. First, observe that if either player arbitrarily decides to implement a deterministic sequence of actions, say $\{W, D, D, W, D, W, \dots\}$, where “W” corresponds to “wait” and “D” corresponds to “dial”, then the other player could have also selected the same sequence $\{W, D, D, W, D, W, \dots\}$, which would result in the connection never being reestablished. While it seems unlikely that Al and Betty would make the same arbitrary choice, this is an example of the worst case outcome, and, without knowing the mechanism by which the other agent will select actions, it seems that we are stuck with this worst case as the only way to evaluate arbitrary decisions.

In contrast, randomized strategies make a lot of sense. For example, if Al implements a strategy of dialing with probability $p \in [0, 1]$ at each dial-wait opportunity, independently of his or Betty's actions at previous stages, and if Betty similarly dials with probability $q \in [0, 1]$ at each stage, then the coordination process reduces to a Markov chain, and the expected number of stages until the connection is reestablished works out to be $[p(1 - q) + (1 - p)q]^{-1}$, plotted in Figure 1. Note that if Al and Betty choose $p = 0$ and $q = 1$, respectively, then they reconnect in one stage, the best possible outcome. What complicates matters is that another solution achieves the same result, namely $p = 1$ and $q = 0$, and if Al and Betty can't agree on which one of these two solutions to pick, then they wind up achieving the worst possible outcome, either $(p, q) = (0, 0)$ or $(p, q) = (1, 1)$, for which the connection is never reestablished.

Analyzing this example in game-theoretic terms, we observe that both of the best-case solutions $(p, q) = (1, 0)$ and $(p, q) = (0, 1)$ are Nash equilibria in the common interest game defined by the cost (disutility) function $f(p, q) = [p(1 - q) + (1 - p)q]^{-1}$. (Neither player can deviate unilaterally to achieve a lower value of cost.) There is a third Nash equilibrium solution where each player dials with probability one half, so that (regardless of what action the other player chooses) the per-stage probability of reestablishing the connection is one-half, and the expected number of stages to terminate the process is two. Mathematically, $f(.5, q) = 2 = f(p, .5)$ for all $p, q \in [0, 1]$, which also shows that $(p, q) = (.5, .5)$ is a saddle-point (minimax) solution for the game.

Considering that there are multiple Nash equilibrium solutions, game theory doesn't offer much insight. Given that Al and Betty will only play actions that are consistent with Nash equilibrium solutions, then which equilibrium should they choose? Should they choose one of the Nash equilibria with lowest expected cost? If so, which one? (Again, if Al and Betty disagree on this point then they experience the worst case outcome.) Should Al and Betty settle on the mixed strategy equilibrium $(p, q) = (.5, .5)$? In what sense would this be rational? Is there something more to the solution $(p, q) = (.5, .5)$ than the fact that it is a Nash equilibrium in the common interest game defined by $f(p, q)$? (If so, then the problem goes beyond one of

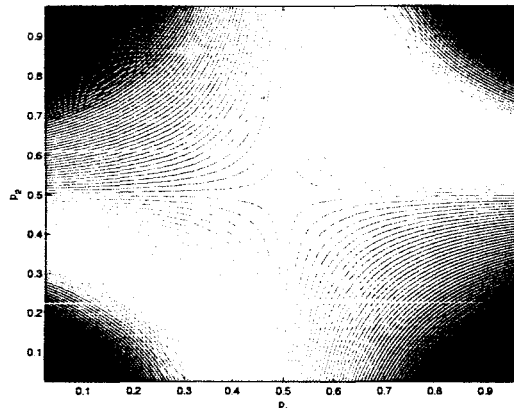


Figure 1: Expected number of stages to reestablish connection (note log-scale on the z-axis), where p is the per-stage probability that Al dials and q is the per-stage probability that Betty dials

“equilibrium selection.”)

Concepts from game theory are generally not strongly useful for distributed decision-making without coordination. Notions of equilibria in games are founded on the idea that the game will be played many times, with the players gaining some information each time. Under such an assumption, it is not critical to define a computational procedure for finding equilibrium policies. Indeed, much of the literature in game theory focuses on defining concepts of equilibria and establishing the conditions under which they exist. In our problem, however, we must adopt an algorithmic approach, defining a reasonable concept of a solution to the problem and providing an algorithm by which this solution can be found. In this sense, we propose to treat distributed decision-making without coordination much more like a decision theoretic or optimization problem than a formal game.

3 General Coordination Processes

We generalize the dial-wait example by defining the notion of *coordination processes* as follows. A coordination process is a discrete-stage dynamic process, similar to a Markov control process [8] and the dynamic team model of [10], whose *state* is denoted by $x \in X$, where X is a Borel space representing the set of all possible operating conditions (the *state space*) associated with the coordination process. For each state $x \in X$ there is an associated set $U(x) = \{1, 2, \dots, N_x\}$ of *actors* whose actions jointly determine the outcome of the process per stage, where N_x is the (possibly infinite) number of actors associated with state $x \in X$. Each actor $i \in U(x)$ has an associated set of actions (*action set*) $A_i(x)$, where $A_i(x)$ is a Borel space. A *feasible action profile* $\alpha = (a_1, \dots, a_{N_x}) \in \prod_{i=1}^{N_x} A_i(x) \equiv A(x)$, represents a collection of actions selected independently by the actors $U(x)$ when the coordination process is in state $x \in X$. We assume

that the set $K = \{(x, \alpha) \mid x \in X, \alpha \in A(x)\}$ is a Borel measurable subset of $X \times A$. In jointly (and independently) selecting a feasible action profile $\alpha \in A(x)$ the actors collectively determine probability distribution for the next state of the coordination process. Let \mathcal{Q} denote the stochastic kernel on X given K that serves as the *state transition law* of the coordination process. The actors' joint selection of actions in a feasible action profiles also determines the probability distribution $F_C^{(x, \alpha)}$ for the random system-level cost associated with the state transition, where $(x, \alpha) \in K$. For convenience, let $c_x(\alpha)$ denote the expected cost associated with state transitions from $x \in X$ under the feasible action profile $\alpha \in A(x)$. We assume that all actors have perfect knowledge of the state transition law \mathcal{Q} , whereas, depending upon the application, we may or may not assume that the transition cost distribution function is known.

Coordination process evolves in discrete time, starting from an initial state $x^0 \in X$. Let $H^0 = \{x^0\}$ denote the initial history of the process. Each actor $i \in U(x^0)$ observes a (possibly set-valued) function of the initial history $\zeta_i^0(H^0)$, and upon making this observation independently chooses an action $a_i^0 \in A_i(x^0)$. Let $\alpha^0 = (a_1^0, \dots, a_{N_{x^0}}^0) \in A(x^0)$ represent aggregation of all actions selected by the actors in $U(x^0)$. Then, according to the state transition law \mathcal{Q} and the transition-cost distribution function $F_C^{(x^0, \alpha^0)}$, the coordination process transitions to a new state $x^1 \in X$ and experiences a transition cost c^0 . Let $H^1 = \{x^0, \alpha^0, c^0, x^1\}$ denote the history of the process at stage 1. The process continued similarly over some (possibly random and/or infinite) time horizon T . That is, at stage $t = 1, \dots, T - 1$:

1. Given the history H^t , each actor $i \in U(x^t)$ observes a (possibly set-valued) function of the initial history $\zeta_i^t(H^t)$, and independently chooses an action $a_i^t \in A_i(x^t)$.
2. Under the feasible profile $\alpha^t = (a_1^t, \dots, a_{N_{x^t}}^t) \in A(x^t)$, the system transitions to the next state $x^{t+1} \in X$ and experiences a transition cost c^t according to the state transition law \mathcal{Q} and the transition-cost distribution function $F_C^{(x^t, \alpha^t)}$, with $H^{t+1} = H^t \cup \{\alpha^t, c^t, x^{t+1}\}$.

Example 1 (Dial-Wait Revisited) *The Dial-Wait example can be seen to be a coordination process in which there are two states $X = \{1, \Omega\}$, with two actors per state $U(1) = U(\Omega) = \{1, 2\}$, each actor being one of the two parties in the call. State $x = 1$ corresponds to the “disconnected” state, where the two callers are still trying reestablish their connection, and state $x = \Omega$ corresponds to “connected” state, where the two callers have finally managed to establish their connection. While in state $x = 1$ both actors $i = 1, 2$ have two pure actions available $\tilde{A}_i(1) = \{1, 2\}$, where $\tilde{a}_i = 1$ corresponds to the decision to dial and $\tilde{a}_i = 2$ corresponds to the decision to wait. Allowing both actors to randomize their decisions we have $A_i(1) = [0, 1]$, where a_i corresponds to the probability that $\tilde{a}_i = 1$, $i = 1, 2$, and the probability of transitioning from $x = 1$ to Ω is then $a_1(1 - a_2) + (1 - a_1)a_2$. Since both actors seek to be reconnected as quickly as possible, they perceive (deterministically) a system-level transition cost $c(1) = 1$ for all state transitions from $x = 1$, including self-transitions. The coordination process terminates as soon as it transitions into $x = \Omega$, so that the total cost associated with reconnecting is $\sum_{t=0}^{T-1} c(1) = T$, where T is the random number of stages of dialing/waiting until reconnecting.*

While the Dial-Wait example is very simple, the mathematical framework for coordination processes above is quite general, having much of the same structure as that for Markovian control processes [6, 4, 2, 3, 8, 9]. Of course, coordination processes are more general since they allow for multiple decision makers (i.e. actors) who interact through their actions as time evolves. What distinguishes our framework from earlier work on dynamic noncooperative games [1] is that we make the extreme simplifying assumption that all players have the *same* objective: to minimize a single system-level notion of cost. It is this special structure that provides the opportunity for advances above and beyond the confines of game theory as known today, and because of this we refer to the decision makers in our model as “actors” rather than as players.

Despite assuming that all actors have the same preference structure, the general class of coordination processes above is too unstructured to allow for much progress (either analytically or computationally). Fortunately, at least for the applications that we have in mind, it is possible to identify a simpler model whose structure can be exploited in defining new solution concepts and in characterizing and computing optimal coordination strategies. In robotic applications, for example, it makes sense to assume that the set of actors $U(x)$ is fixed, i.e. not dependent on the state of the process. Also, for the planning-type applications, it makes sense to assume that the state space X and the sets of pure actions available to each actor are all finite. Thus, we propose to focus attention on coordination processes that satisfy the following structural assumption:

Assumption 1 *The following are true.*

1. *The state space X is finite.*
2. *The same set of actors applies at all states, and this set is finite. That is, $U(x) = U \equiv \{1, 2, \dots, N\}$ for all $x \in X$, where N denotes the (constant) number of actors.*
3. *Each actor $i \in U$ has a finite set of pure actions $\tilde{A}_i(x)$ and chooses mixed actions over $\tilde{A}_i(x)$ as the coordination process evolves, i.e.*

$$A_i(x) = \left\{ \mathbf{a}_i = (a_{i,j})_{j \in \tilde{A}_i(x)} \mid a_{i,j} \geq 0, j \in \tilde{A}_i(x), \text{ and } \sum_{j \in \tilde{A}_i(x)} a_{i,j} = 1 \right\}.$$

The coordination process evolves as a sequence of each actor i choosing “mixed actions” $\mathbf{a}_i \in A_i(x)$, where the decision at each stage is what probability distribution over \tilde{A}_i should be played.

4. *For each pair of states x and \bar{x} in X and for each profile of pure actions $\tilde{\alpha} = (\tilde{\alpha}_1, \dots, \tilde{\alpha}_N) \in \prod_{i=1}^N \tilde{A}_i(x)$, there is a corresponding state transition probability $\tilde{p}_{x\bar{x}}(\tilde{\alpha})$, representing the probability of transitioning from x to \bar{x} under the pure actions $\tilde{\alpha}$. Let $p_{x\bar{x}}(\alpha)$ denote the resulting probability of transitioning from x to \bar{x} under the profile of (mixed) actions $\alpha \in A$.*

Note that Assumption 1 essentially restricts attention to coordination processes that have same basic structure as Markov decision processes [11] and (zero-sum) competitive Markovian decision

processes [7]. We plan to focus attention on a class that we term *transient* coordination processes. These are coordination processes in which each actor must identify an appropriate (mixed) action for each state of the process so as to reach an absorbing and cost-free state along a minimum cost path, as in the Dial-Wait example. Problems of this type have very much the character of so-called transient competitive Markovian decision processes [7], although, because of our assumption about a common objective function, we find it convenient to introduce a new solution concept, namely minimum unambiguous value (MUV) solutions, to address the inadequacy of Nash equilibria as observed in the Dial-Wait example.

4 Transient Coordination Processes

In this section, we focus on coordination processes with the property that all actors seek to drive an underlying system to an absorbing, zero-cost state Ω along a minimum cost trajectory through the state space.

Definition 1 *A transient coordination process is a coordination process in which all actors seek to minimize the expected discounted cost associated with the evolution of the system,*

$$\mathbb{E} \left\{ \limsup_{T \rightarrow \infty} \sum_{t=0}^T \gamma^t c^t \mid \text{actor decisions} \right\}, \quad (1)$$

where $\gamma \in [0, 1]$ is a discount factor, subject to Assumption 1 and the following additional assumptions.

1. The coordination process has an absorbing, zero-cost state $\Omega \in X$. That is, there exists a state Ω such that $p_{\Omega\Omega}(\alpha) = 1$ and $c_{\Omega}(\alpha) = 0$ for all $\alpha \in A(\Omega)$.
2. Each actor has perfect knowledge of the distribution function $F_C^{(x,\alpha)}$.
3. In selecting an action at each stage, each actor has knowledge only of (or restricts attention to) the current state of the system x^t at each stage of the process, i.e. $\zeta_i^t(H^t) = x^t$ for all $t = 0, 1, \dots$

Some examples of transient coordination processes follow.

Static Coordination Problems Consider the situation where a given set of actors U are engaged in a single (aggregate) decision making over a set A of feasible action profiles, where the outcome of the process is a random cost C whose distribution function F_C^a is determined by the feasible action profile $a \in A$ selected by the actors. Static coordination problems of this type can be interpreted as transient coordination processes involving two states $X = \{1, \Omega\}$, where (i) the system starts out in state $x = 1$ (with $A(1) = A$ and $F_C^{(1,\alpha)} = F_C^a$) and (ii) the system transitions immediately to the terminal state Ω , which is absorbing and cost free.

Finite-Horizon Coordination Processes Finite-state, finite-horizon processes can similarly be expressed as transient coordination processes. Suppose a given set of actors U make decisions over a predetermined finite time horizon T , where \tilde{X}_n denotes the set of all possible states of the coordination process at stage $n = 0, 1, \dots, T-1$. In this case, the set $X = \tilde{X}_0 \cup \tilde{X}_1 \cup \dots \cup \tilde{X}_{T-1} \cup \{\Omega\}$ can be interpreted as the state space of an equivalent transient coordination process, where the transition probability matrix is such that only transitions from X_n to X_{n+1} for $n = 0, 1, \dots, T-2$ and from X_{T-1} to $\{\Omega\}$ have nonzero probability.

Discounted Cost Coordination Processes Infinite-horizon, finite-state processes with a discounted cost criterion can also be expressed as transient coordination processes. Suppose a given set of actors U make decisions on an infinite time horizon, where

1. the finite set \tilde{X} denotes the state space associated with each stage of the process,
2. $p_{x\bar{x}}(\alpha)$ is the probability of transitioning from $x \in \tilde{X}$ to $\bar{x} \in \tilde{X}$ under the feasible action profile α , and
3. all actors seek to minimize the expected discounted cost objective of Equation (1) with $\gamma < 1$.

Using a well-known trick from the theory of Markov decision processes, we can define an equivalent transient coordination process that evolves over the state space $X = \tilde{X} \cup \{\Omega\}$, where Ω is a cost-free absorbing state, by adjusting the transition probability matrix so that $1 - \gamma$ is the probability of transitioning to Ω from any state $x \in X$ and transitions from $x \in X$ to $\bar{x} \in X$ occur with probability $\gamma \cdot p_{x\bar{x}}(\alpha)$. In this way, the effect of the discount factor shows up as the per-stage probability of not terminating in the equivalent transient model.

Stochastic Shortest Path Coordination Processes A quite general class of transient coordination processes are defined by the undiscounted stochastic shortest path assumptions of [5]. In this case, one can refine Definition 1 by assuming (additionally) that (i) at least one actor $u \in U$ has the ability to guarantee termination of the process (i.e. that the system will eventually transition to Ω) with probability one and (ii) whenever the actors behave in such a way that there is a chance that the system never terminates, then the expected total cost of the process is infinite.

In the research under this grant, we have restricted attention to static coordination problems. A proposed solution concept for this class problems is described below.

5 Minimum Unambiguous Value

The central issue in defining an appropriate solution concept is considering how to deal with the fact that if the agents are faced with a set of multiple (mixed) action profiles that are equally attractive in terms of cost, a failure of coordination may lead to an expected cost value that is

worse than that of the profiles in the set. Below we propose a solution concept that is grounded in the idea that any optimal action profile for a coordination problem should be unambiguous in the sense of not being subject to degradation if any of the agents elect to take an alternative action that is equally attractive in terms of cost.

To formalize this notion, we first define some notation. Given two mixed action profiles, $\xi^0 = (x_1^0, x_2^0, \dots, x_N^0)$ and $\xi^1 = (x_1^1, x_2^1, \dots, x_N^1)$, the worst case confusion that can arise between ξ^0 and ξ^1 is

$$\phi(\xi^0, \xi^1) = \max_{(j_1, j_2, \dots, j_N) \in \{0,1\}^N} f(x_1^{j_1}, x_2^{j_2}, \dots, x_N^{j_N}) \quad (2)$$

Given a mixed action profile $\xi = (x_1, x_2, \dots, x_N)$, define¹

$$\Phi(\xi) = \sup_{\xi^0 : f(\xi^0) = f(\xi)} \phi(\xi^0, \xi). \quad (3)$$

Note that $\Phi(\xi) \geq f(\xi)$ for all ξ since $\phi(\xi, \xi) = f(\xi)$. We say that a mixed action profile ξ has an *ambiguous value ceiling* if $\Phi(\xi) = f(\xi)$. In other words, if ξ has an ambiguous value ceiling, then there is no risk that a failure to coordinate in choice of equally attractive profiles will lead to an objective value worse than that of ξ^1 . If ξ does not have an ambiguous value ceiling, i.e. if $\Phi(\xi) > f(\xi)$, then we say that ξ is *confusable*.

In terms of the above definition, we assert that one can cast the problem faced by the agents as being one of finding a solution that has smallest expected cost among all solutions that have an ambiguous value ceiling, a point which we call a solution of *minimum unambiguous value*, or MUV for short.

Definition 2 (MUV) We say that an action profile ξ^1 is MUV if (i) ξ^1 has an ambiguous value ceiling and (ii) all ξ such that $f(\xi) < f(\xi^1)$ are confusable.

Example 2 (Dial-Wait Revisited) Referring back to the Dial-Wait example of Section 2, we observe that three mixed strategy profiles have an ambiguous value ceiling:

1. The solution where both Al and Betty dial with probability one, i.e. $x_{1,dial} = p_1 = 1$ and $x_{2,dial} = p_2 = 1$, has an ambiguous value ceiling of one.
2. The solution where both Al and Betty wait with probability zero, i.e. $x_{1,dial} = p_1 = 0$ and $x_{2,dial} = p_2 = 0$, also has an ambiguous value ceiling of one.
3. The solution where both Al and Betty dial independently with probability one half, i.e. $x_{1,dial} = p_1 = .5$ and $x_{2,dial} = p_2 = .5$, has ambiguous value ceiling of one half.

The last solution (i.e. $x_{1,dial} = p_1 = .5$ and $x_{2,dial} = p_2 = .5$) is such that any attempt to achieve a lower expected value, namely lower than one half, can be confused with another solution achieving the same value resulting in greater expected cost. Thus, $(p_1, p_2) = (.5, .5)$ is the MUV solution to the Dial-Wait problem.

¹In Section 7.3 we propose an alternative definition in which the supremum is over all ξ^0 such that $f(\xi^0) \leq f(\xi)$.

In general, if a unique MUV solution ξ exists, then we may take ξ as the solution to the identical interest game defined by f . Indeed, the MUV solution concept seems to adequately capture the essence of rational decision making under the axiom of no arbitrary action. However, there are still many aspects of the concept that need to be addressed.

Open Problem 1 *If ξ^1 is MUV, and if ξ is such that $f(\xi) < f(\xi^1)$, is it true that $\Phi(\xi) > f(\xi^1)$?*

In addition, we need to understand the conditions under which MUV solutions exist. If existence is not generally guaranteed (as it may not be, since the sets of solutions that have ambiguous value ceilings may not have nice topological properties), then it may be useful to consider notions of approximate (or ϵ -) MUV.

Open Problem 2 *Under what conditions will a MUV or ϵ -MUV solution exist?*

We also need to understand the relationships, if any, between MUV solutions and the various notions of equilibria that have been defined for games and other related problems. Are the agents facing a problem of equilibrium selection, or something quite different that will require the development of a new algorithm? The latter case is more likely and it exposes the need to understand the computational complexity of the problem. This is a non-trivial task, since it is not apparent how to formulate the task of finding a MUV point as a standard computational or optimization problem. It is likely that the problem is NP-hard and we will need to turn to the development of reasonable heuristics or approximation algorithms. This would be an acceptable outcome; indeed our primary purpose in studying the theory of transient coordination processes is to develop insight that can be the basis for the design of effective heuristics for real-world applications.

5.1 Alternative Characterizations of MUV

In this section we consider some alternative characterizations of MUV that may prove useful in defining a computational approach to solving static coordination processes.

5.1.1 Alternative 1

For an alternative characterization of MUV, it is convenient to define

$$v^* = \inf V \triangleq \left\{ v \mid \begin{array}{l} \text{there exists } \xi^1 \text{ such that} \\ 1. f(\xi^1) = v \text{ and} \\ 2. \Phi(\xi^1) = f(\xi^1) \end{array} \right\}. \quad (4)$$

Note that the set $V \subset \Re$ is non-empty. For example, it contains $v_{\max} = \max_{a_1, a_2, \dots, a_N} f(a_1, a_2, \dots, a_N)$.

Open Problem 3 *How do we characterize V ? Is it a finite set? Does it have a minimum value?*

To establish a connection with Definition 2, we observe the following.

Lemma 1 *A mixed strategy profile ξ is MUV if and only if*

$$f(\xi) = v^* \quad \text{and} \quad \Phi(\xi) = v^*,$$

with v^ as defined in Equation (4).*

Note that if ξ is not the unique MUV solution, then we have the following.

Lemma 2 *If ξ^1 and ξ^2 are both MUV, then $\phi(\xi^1, \xi^2) \leq v^*$*

In other words, if ξ^1 and ξ^2 are both MUV (achieving expected cost v^*), then any profile ξ^3 constructed as a permutation of ξ^1 and ξ^2 cannot have cost greater than v^* .

Open Problem 4 *Can the inequality in Lemma 2 be strict?*

Open Problem 5 *Given two MUV solutions $\xi^1 = (x_1^1, x_2^1, \dots, x_N^1)$ and $\xi^2 = (x_1^2, x_2^2, \dots, x_N^2)$, and defining $\xi^3 = (x_1^3, x_2^3, \dots, x_N^3) = (x_1^{k_1}, x_2^{k_2}, \dots, x_N^{k_N})$ with arbitrary permutation $(k_1, k_2, \dots, k_N) \in \{1, 2\}^N$, is it true in general that $\phi(\xi^0, \xi^3) \leq v^*$ for all ξ^0 such that $f(\xi^0) = v^*$.*

If the the answer to the open problem above is “yes” for all N , then ξ^3 is such that confusing it with any other profile ξ^0 such that $f(\xi^0) = v^*$ results in cost less than or equal to v^* .

Open Problem 6 *Suppose ξ^1 is such that $\Phi(\xi^1) = f(\xi^1)$ and suppose ξ^2 is such that $\Phi(\xi^2) = f(\xi^2) < f(\xi^1)$, is it true that $\phi(\xi^1, \xi^2) \leq f(\xi^1)$?*

5.1.2 Alternative 2

Given $v \in \mathfrak{R}$, define

$$\Psi(v) = \inf_{\xi: f(\xi)=v} \Phi(\xi). \quad (5)$$

Now define

$$\tilde{v}^* = \inf \tilde{V} \triangleq \{ v \mid v = \Psi(v) \}. \quad (6)$$

Conjecture 1 *It is true that $\tilde{v}^* = v^*$, in which case MUV corresponds to a minimal fixed point of Ψ .*

6 Evaluating MUV Solutions

The concepts that we have considered up to this point do not by themselves suggest a method for actually computing a MUV solution. In general this seems to be a very hard problem, and well outside the scope of this initial research effort. Some special cases, however, are tractable as discussed below.

6.1 Two-Player, Two-Action Games

Consider the static coordination problem defined by the matrix

$$F = \begin{bmatrix} f(a_{1,1}, a_{2,1}) & f(a_{1,1}, a_{2,2}) \\ f(a_{1,2}, a_{2,1}) & f(a_{1,2}, a_{2,2}) \end{bmatrix},$$

where the set of pure actions available to Player 1 is $A_1 = \{a_{1,1}, a_{1,2}\}$ and the set of pure actions available to Player 2 is $A_2 = \{a_{2,1}, a_{2,2}\}$. To simplify the notation a bit, we will use f_{ij} to refer to $f(a_{1,i}, a_{2,j})$, the cost of the action profile $a_{1,i}, a_{2,j}$, for $i, j \in \{1, 2\}$. We parameterize Player 1's mixed strategies as $x_1 = (p_1, (1 - p_1))$ and Player 2's strategies as $x_2 = (p_2, (1 - p_2))$, where p_1 is the probability that Player 1 chooses $a_{1,1} \in A_1$ and p_2 is the probability that Player 2 chooses $a_{2,1} \in A_2$. With this parameterization in mind, we can express $f(x_1, x_2)$ as

$$V(p_1, p_2) = p_1 p_2 f_{11} + (1 - p_1) p_2 f_{21} + p_1 (1 - p_2) f_{12} + (1 - p_1) (1 - p_2) f_{22}. \quad (7)$$

To make this problem interesting, let us assume that $f_{11} = k = f_{22}$ and $f_{12}, f_{21} \geq k$. If the former inequalities are strict, then there are two conflicting optimal solutions to the problem, and it becomes interesting to consider the MUV solution to the problem.

Lemma 3 *Assuming that $f_{11} = k = f_{22}$ and $f_{12}, f_{21} \geq k$, with $f_{12} + f_{21} > 2k$, the mixed action pair parameterized by*

$$p_1^* = \frac{(f_{21} - k)}{(f_{12} + f_{21} - 2k)} \quad p_2^* = \frac{(f_{12} - k)}{(f_{12} + f_{21} - 2k)} \quad (8)$$

is a MUV solution to the static coordination game defined by F , and

$$V(p_1^*, p_2^*) = k + \frac{(f_{12} - k)(f_{21} - k)}{(f_{12} + f_{21} - 2k)}. \quad (9)$$

Proof: By hypothesis, we have that

$$V(p_1, p_2) = k + (f_{12} - k)p_1 - (f_{12} + f_{21} - 2k)p_1 p_2 + (f_{21} - k)p_2.$$

Plugging in p_2^* we have

$$\begin{aligned} V(p_1, p_2^*) &= k + [(f_{12} - k) - (f_{12} + f_{21} - 2k)p_2^*]p_1 + (f_{21} - k)p_2^* \\ &= k + (f_{21} - k)p_2^* \\ &= k + \frac{(f_{12} - k)(f_{21} - k)}{(f_{12} + f_{21} - 2k)} \end{aligned}$$

for all $p_1 \in [0, 1]$. Similarly, given that Player 1 adopts the mixed strategy p_1^* , we have

$$V(p_1^*, p_2) = k + \frac{(f_{12} - k)(f_{21} - k)}{(f_{12} + f_{21} - 2k)}$$

for all $p_2 \in [0, 1]$. Thus, the mixed action pair parameterized by p_1^* and p_2^* has an ambiguous value ceiling. Moreover, it can be verified that any p_1, p_2 such that $V(p_1, p_2) < V(p_1^*, p_2^*)$ is confusable. Q.E.D.

6.2 Symmetric Problems

For static coordination problems that are symmetric in the sense that

$$f(x_1, x_2, \dots, x_N) = f(x_{p(1)}, x_{p(2)}, \dots, x_{p(N)}) \quad (10)$$

for any permutation $\{p(1), p(2), \dots, p(N)\}$ of $\{1, 2, \dots, N\}$, it is possible to infer a property of the MUV solution.

Conjecture 2 *The MUV solution to the symmetric problem is (x^*, x^*, \dots, x^*) , where $x^* = \arg \min f(x, x, \dots, x)$.*

7 Alternative Solution Concepts

In this section we collect a number of solution concepts whose relationship to MUV have yet to be determined precisely.

7.1 MUV+

In case the answer to Open Problem 1 turns out to be “No,” then we could consider an alternative version of the MUV solution concept, as follows.

Definition 3 (MUV+) *We say that an action profile ξ^1 is MUV+ if (i) ξ^1 has an ambiguous value ceiling and (ii) all ξ such that $f(\xi) < f(\xi^1)$ are such that $\Phi(\xi) > f(\xi^1)$.*

Open Problem 7 *The existence, uniqueness, implications of non-uniqueness, and characterization of MUV+ solutions all need to be established.*

7.2 Minimax Confusion

If we are willing to give up on the requirement that a “solution” to the identical interest game must have an ambiguous value ceiling, then we may consider yet another alternative to MUV, as follows. Define

$$M^* = \inf_{\xi^1 = (x_1^1, x_2^1, \dots, x_N^1)} \Phi(\xi^1) \quad (11)$$

Definition 4 *If ξ^1 achieves the infimum in Equation (11), then we say it minimizes confusion in the identical interest game defined by f .*

Open Problem 8 *The existence, uniqueness, implications of non-uniqueness, and characterization of minimax confusion solutions all need to be established.*

7.3 Revised notion of “ambiguous value ceiling”

Define

$$\tilde{\Phi}(\xi) = \sup_{\xi^0 : f(\xi^0) \leq f(\xi)} \phi(\xi^0, \xi), \quad (12)$$

and replace Φ with $\tilde{\Phi}$ in all earlier definitions, particularly in the definition of “ambiguous value ceiling.” By expanding the domain of the supremum in the definition, we do not significantly change our internalization of the MUV concept. However, considering the final remark of Open Problem 6, this may actually be the best definition.

References

- [1] T. Basar and G. J. Olsder. *Dynamic Noncooperative Game Theory*, 2nd ed. Academic Press, London, 1995.
- [2] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, Belmont, MA, 1995.
- [3] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, Belmont, MA, 1995.
- [4] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete Time Case*. Academic Press, New York, 1978.
- [5] D. P. Bertsekas and J. N. Tsitsiklis. Analysis of Stochastic Shortest Path Problems. *Mathematics of Operations Research*, 16(3):580–595, 1991.
- [6] E. B. Dynkin and A. A. Yushkevich. *Controlled Markov Processes*. Springer-Verlag, Berlin, 1975.
- [7] J. A. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer, New York, 1997.
- [8] O. Hernandez-Lerma and J. B. Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer-Verlag, New York, 1996.
- [9] O. Hernández-Lerma and J. B. Lasserre. *Further Topics on Discrete-Time Markov Control Processes*. Springer-Verlag, New York, 1999.
- [10] J. Marschak and R. Radner. *Economic Theory of Teams*. Yale University Press, New Haven, 1972.
- [11] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York, 1994.